different online discussions or even different web sites from the answer that is being scored. The set of supporting evidence is stored in supporting evidence data store **920**.

[0055] At step **925**, the process generates a supporting evidence score based on factors of the supporting evidences such as a quantity of supporting evidence for the question that is being scored as well as the quality of supporting evidence that has been found supporting the question. The supporting evidence score is stored in supporting evidence score data store **930**.

[0056] At step **935**, the process identifies one or more "ratings" of the answer post, if any, that were included in the online threaded discussion. As previously mentioned, some online threaded discussions provide for a 'rating' that is based on the quality of the answer and/or a rating of the online threaded discussion. This answer post score is stored in post score data store **940**.

[0057] At step **945**, the process identifies an experience or expertise level of the individual that provided the answer post based on possible labels attached to the individual poster (e.g., "expert," "experienced, etc.) a proliferation level found for the individual poster, and other experience factors found in the supporting evidence for the individual poster's expertise pertaining to the question and the answer included in the post. The post provider score that is generated is stored in post provider score data store **950**.

[0058] At step **955**, the process identifies follow-up comments from other posters that provide indications regarding the answer posts correctness or accuracy. The follow-up comments that are identified are used to generate a follow-up score that is stored in follow-up score data store **960**.

[0059] At step **965**, the process identifies any other reliability factors that might be present or inherent in the online threaded discussion and/or the answer post and a score is generated based on such other factors. The score is stored in other factors score data store **970**.

[0060] At step **980**, the process calculates an overall score based on the component scores (supporting evidence score **930**, answer post score **940**, post provider score **950**, follow-up score **960**, and other factors score **970**). The overall score that is calculated is associated with the answer post and added to candidate answers data store **370**. In one embodiment, when all of the answers have been scored, the candidate answer with the best, or highest, score is selected as being the most likely answer to the user's question. Processing then returns to the calling routine (see FIG. **8**) at **995**.

[0061] The flowchart and block diagrams in the Figures illustrate the architecture, functionality, and operation of possible implementations of systems, methods and computer program products according to various embodiments of the present invention. In this regard, each block in the flowchart or block diagrams may represent a module, segment, or portion of code, which comprises one or more executable instructions for implementing the specified logical function (s). It should also be noted that, in some alternative implementations, the functions noted in the block may occur out of the order noted in the figures. For example, two blocks shown in succession may, in fact, be executed substantially concurrently, or the blocks may sometimes be executed in the reverse order, depending upon the functionality involved. It will also be noted that each block of the block diagrams and/or flowchart illustration, and combinations of blocks in the block diagrams and/or flowchart illustration, can be implemented by special purpose hardware-based systems that perform the specified functions or acts, or combinations of special purpose hardware and computer instructions.

[0062] While particular embodiments of the present invention have been shown and described, it will be obvious to those skilled in the art that, based upon the teachings herein, that changes and modifications may be made without departing from this invention and its broader aspects. Therefore, the appended claims are to encompass within their scope all such changes and modifications as are within the true spirit and scope of this invention. Furthermore, it is to be understood that the invention is solely defined by the appended claims. It will be understood by those with skill in the art that if a specific number of an introduced claim element is intended, such intent will be explicitly recited in the claim, and in the absence of such recitation no such limitation is present. For non-limiting example, as an aid to understanding, the following appended claims contain usage of the introductory phrases "at least one" and "one or more" to introduce claim elements. However, the use of such phrases should not be construed to imply that the introduction of a claim element by the indefinite articles "a" or "an" limits any particular claim containing such introduced claim element to inventions containing only one such element, even when the same claim includes the introductory phrases "one or more" or "at least one" and indefinite articles such as "a" or "an"; the same holds true for the use in the claims of definite articles.

What is claimed is:

1. A method, in an information handling system comprising a processor and a memory, of mining threaded online discussions, the method comprising:

performing, by the information handling system, a natural language processing (NLP) analysis of one or more threaded discussions pertaining to a given topic, wherein the analysis is performed across one or more web sites with each of the web sites including one or more of the threaded discussions, wherein the analysis results in a plurality of harvested discussions;

identifying a question from the harvested discussions;

identifying a plurality of candidate answers from the harvested discussions, wherein each of the plurality of candidate answers pertain to the identified question;

aggregating and merging a selected plurality of harvested discussions corresponding to each of the candidate answers, wherein the selected plurality of harvested discussions are supporting evidence corresponding to the respective candidate answer;

generating a supporting evidence score based on one or more factors of the supporting evidence for each of the candidate answers; and

scoring each of the plurality of candidate answers, wherein the scoring calculates an overall score corresponding to each of the candidate answers, wherein the overall score is based upon at least the supporting evidence score.

2. The method of claim **1** further comprising:

comparing a plurality of questions found in the threaded discussions to a posed question, wherein the identified question matches the posed question; and

adding one or more of the correlated harvested discussions to a corpus that is utilized in a deep question answering system.